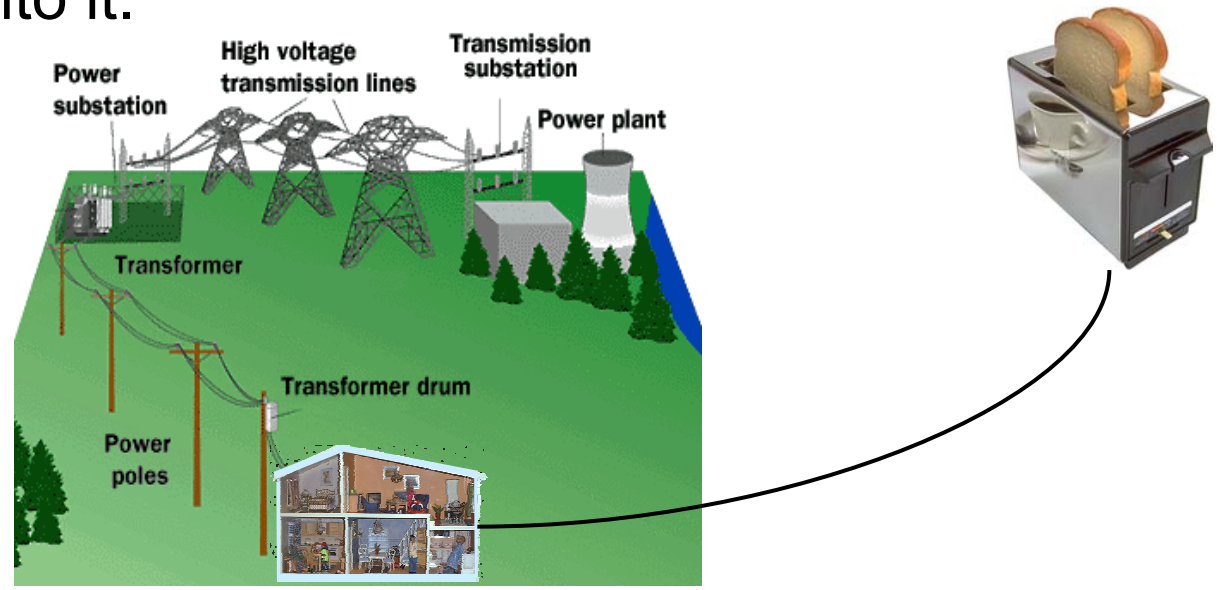
caBIG Applications Overview

Sue Dubman
NCI Center for Bioinformatics

February 19, 2004

The Role of Applications in caBIG

- The original designers of the power grid understood that the value of the grid lay in the appliances that plugged into it.

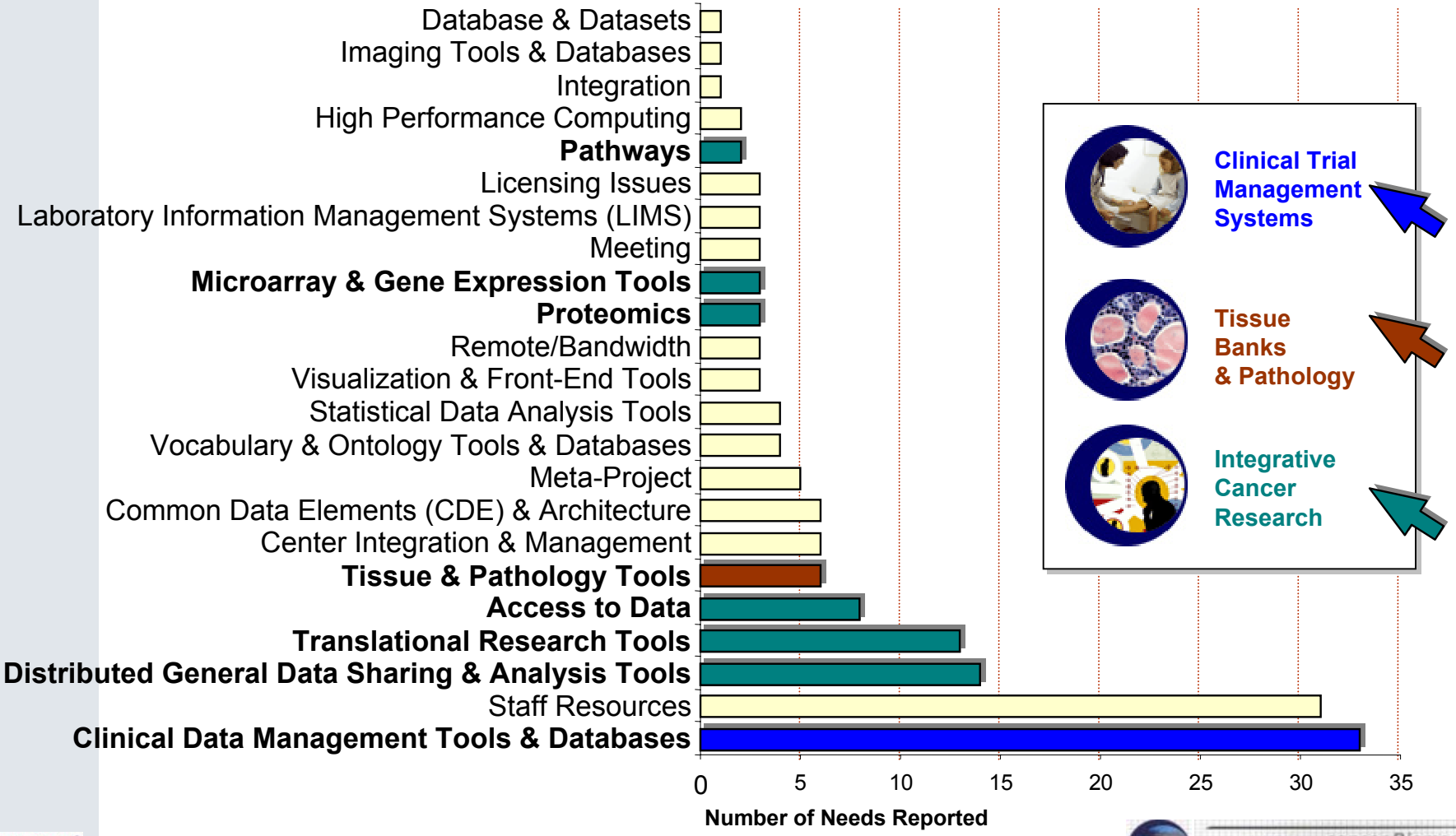


- There is an an analog with caBIG. The applications add value to caBIG and make it worth using.

Application Selection Process

- Key goals and considerations:
  - Ability of potential participants to test the concept and feasibility of the proposed grid
  - Selecting a pilot group that is representative of the varying bioinformatics needs and capabilities
  - Identifying projects that would broadly benefit the Cancer Research Community
  - Get started with already available resources
  - Be as inclusive as possible

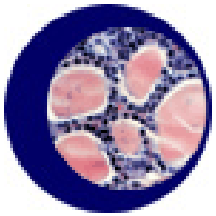
Common needs helped shape priority areas for the caBIG pilot activities



Three Application Workspaces in this Pilot Phase

Clinical Trial Management Systems
Will address the need for consistent, open and comprehensive tools for clinical trials management.

Integrative Cancer Research
Will provide tools and systems to enable integration and sharing of information.



Tissue Banks & Pathology Tools
Will provide for the integration, development, and implementation of tissue and pathology tools.

Clinical Trial Management Systems: Workspace Goals

- Improve clinical research productivity
- Integration of the clinical trials infrastructure with the research infrastructure
- Allow sharing of de-identified data across trials and sites
- Facilitate greater standardization of coding, vocabulary, nomenclature and data collection methods as part of clinical trials
- Define specifications to be followed by developers of new or existing clinical trials systems intended to be compatible with caBIG
- Consistency with NIH Roadmap direction (http://nihroadmap.nih.gov/) .

Clinical Trial Management Systems: Workspace Participants

Developers	Adopters	Working Group
• University of Pittsburgh • City of Hope (Plus-existing caBIG-compatible tools from NCICB)	• Wake Forest University • Duke University • University of California Irvine – Chao Family • Case Western Reserve University – Ireland • University of Pennsylvania – Abramson	• University of California San Francisco • Yale University • Northwestern University – Robert H. Lurie • University of Nebraska – Eppley • University of Iowa – Holden • University of Minnesota • University of Wisconsin • Vanderbilt University – Ingram (Plus – Representatives from each Developer and Adopter site.)

Clinical Trial Management Systems: Workspace Scope

- Near-term:
  - Adoption of existing tools that are already caBIG-compatible or can be quickly made caBIG-compatible
  - Development of a strategic vision for the workspace, detailed specifications, and a defined path for migration
  - Work with Integrated Cancer Research Workspace on clinical trials interoperability
  - Work with vocabulary workspace to facilitate greater standardization of coding, vocabulary, and nomenclature
  - Work with architecture workspace to define specifications for caBIG compatibility, and to define infrastructure requirements
- Ultimately, the efforts should result in a set of related and interoperable tools for clinical data management, arising from the efforts and specific interests of all the Workspace participants

Clinical Trial Management Systems: Example Needs

Example needs within the Clinical Trials Workspace, from August 2003 Cancer Center Cooperative Development Meetings:

- Implement or significantly revise clinical trial management tools (10+ centers)
- Incorporate caBIG interoperability into existing clinical trials management systems
- Protocol authoring with encoding available to the grid
- Structured means to communicate with the FDA, including IND/NDA and AE
- Access to adverse event databases
- Means to maintain patient data security
- Uniform outcomes measures for Quality-of-Life
- Patient recruitment tools
- Document management tools

Integrative Cancer Research: Workspace Goals

- Facilitate the integration of data not only from different centers, but also data of different types
- Enable translational and integrative research
- Provide for the integration of data from “bench to bedside”, “bedside to bench” and from “bench to bench”
- Facilitate greater standardization of coding, vocabulary, nomenclature and data collection methods as part of integrative cancer research
- Define specifications to be followed by developers of new or existing integrative cancer research tools intended to be compatible with caBIG
- Consistency with NIH Roadmap direction (http://nihroadmap.nih.gov/) .

Integrative Cancer Research: Workspace Participants

Developers	Adopters	Working Group
▪ Burnham Institute ▪ Dartmouth – Norris Cotton ▪ Duke University ▪ Georgetown University – Lombardi ▪ University of California San Francisco ▪ University of Chicago ▪ University of Iowa – Holden ▪ University of North Carolina-Lineberger ▪ Cold Spring Harbor ▪ Memorial Sloan-Kettering ▪ Fox Chase ▪ Thomas Jefferson University – Kimmel ▪ Washington University-Siteman (Plus-existing caBIG-compatible tools from NCICB)	▪ New York University ▪ Memorial Sloan-Kettering ▪ Oregon Health and Science University ▪ Wistar ▪ University of South Florida-H. Lee Moffitt ▪ University of Pennsylvania – Abramson	▪ Vanderbilt University – Ingram ▪ Columbia University – Herbert Irving ▪ Meyer L. Prentis-Karmanos - University of Michigan ▪ Northwestern University – Robert H. Lurie (Plus – Representatives from each Developer and Adopter site.)

Integrative Cancer Research: Workspace Scope

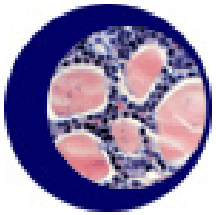
- Extension and implementation of tools to enable researchers to integrate and share data collected from a variety of heterogeneous sources
- caBIG-enable a wide range of tools particularly useful to cancer researchers, with a focus on translational research
- Make data and analytical bioinformatics methods more useful to cancer researchers and data more accessible
- Work with vocabulary workspace to facilitate greater standardization of coding, vocabulary, and nomenclature
- Work with architecture workspace to define specifications for caBIG compatibility, and to define infrastructure requirements

Integrative Cancer Research: Center Needs

Example needs within the Integrative Cancer Research Workspace, from August 2003 Cancer Center Cooperative Development Meetings.

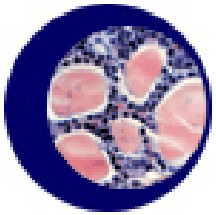
- Need a comprehensive bioinformatics infrastructure to support the integration of clinical and basic science data (10+ Centers)
- Need community-shared platforms for sharing data with collaborators (Multiple Centers)
- Managing large amounts of data challenging
- Better Gene Expression tools
- Want access to large amount of gene expression data, especially in breast cancer
- Better caBIG systems for handling and exchanging pathway data
- Need tools for cross-platform analysis of proteomics data
- Posting of raw data for data mining as opposed to just results (e.g. research data warehouses)
- Need access to data from other Cancer Centers

Tissue Banks and Pathology Tools: Workspace Goals



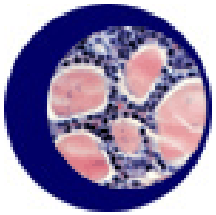
- This Workspace provides for the integration, development, and implementation of tissue and pathology tools to enable sharing of data and information across sites
- These tools will ultimately:
  - Facilitate the integration of data from different studies and from different Centers
  - Provide a platform for rapidly surveying available tissue and pathology resources
  - Develop standard and interoperable tools for working with pathology data and resources

Tissue Banks and Pathology Tools: Workspace Participants



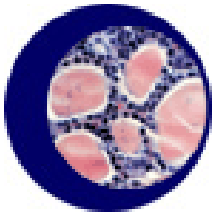
Developers	Adopters	Working Group
• Washington University-Siteman • University of Pittsburgh	• Northwestern University-Robert H. Lurie • University of North Carolina-Lineberger • Dartmouth-Norris Cotton • University of Pennsylvania-Abramson • Thomas Jefferson-Kimmel	• Virginia Commonwealth University – Massey • Jackson Laboratory • Johns Hopkins – Sidney Kimmel • Indiana University • University of Arizona • University of Alabama at Birmingham (Plus – Representatives from each Developer and Adopter site.)

Tissue Banks and Pathology Tools: Workspace Scope



- The integration and implementation of tools for the collection and maintenance of tissue and pathology data
- Develop specifications and interface techniques allowing the creation and federation of repositories to be used by geographically and temporally dispersed clients
- Identify tissue banking and pathology systems already developed by the Centers, and provide means for integration and federation of those resources
- Provide specifications and guidance for the development of future pathology software compatible with the caBIG

Tissue Banks and Pathology Tools: Center Needs



Example needs within the Tissue Banks and Pathology Tools Workspace, from August 2003 Cancer Center Cooperative Development Meetings.

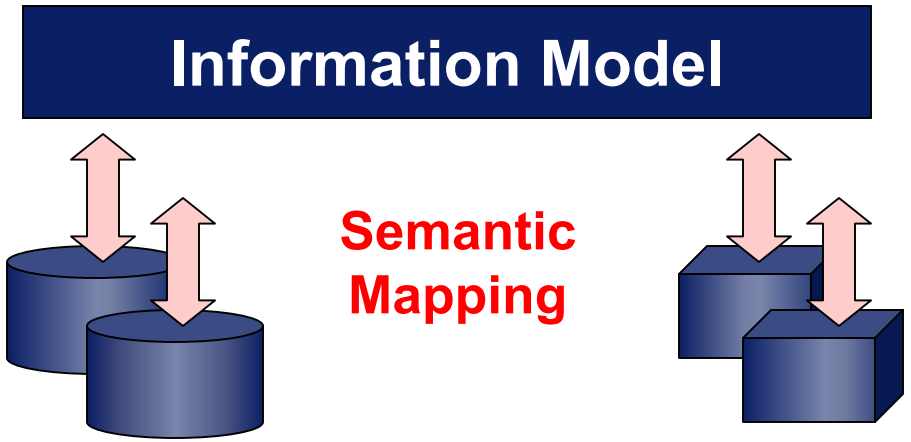
- Support for tissue database work
- Access to tissue banks for research samples
- Sample tracking systems, including bar coding
- Need better community-wide integration of tools for tissue banking database and collaboration
- Tissue pathology bioinformatics

Criteria for Success

- caBIG Infrastructure that supports Application Needs
- caBIG on *SPEED*: Delivering some base capability quickly for those in desperate need of capabilities
- Creating value: Defining long term, strategic vision and start to measurably deliver on those expectations
- Truly open environment that recognizes the diversity that can exist.

Critical Success Factor: caBIG Infrastructure that Supports Application Needs

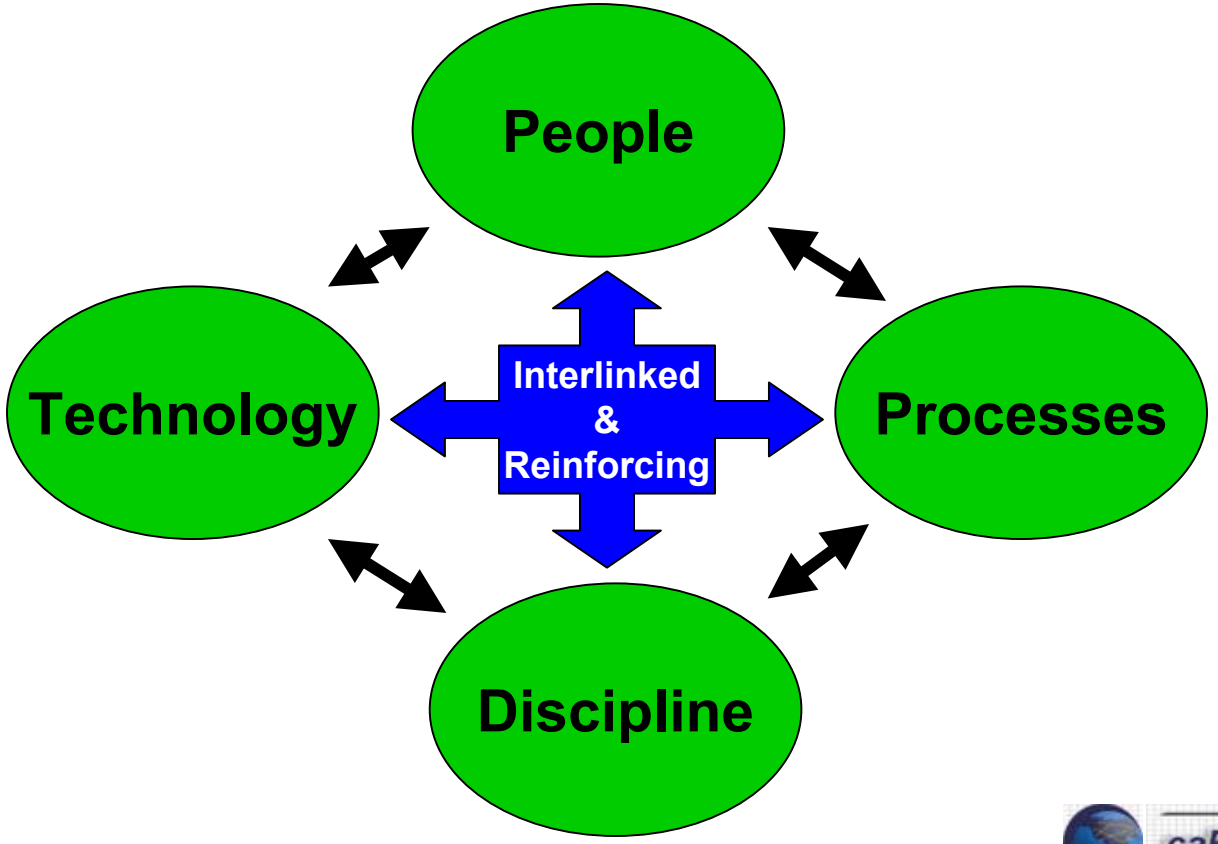
- Transparency
- Pervasiveness
- Semantic Interoperability
- Common components



- Common model across all domains of interest
- Foundation of rigorously defined data types
- Methodology for interfacing with controlled vocabularies

Critical Success Factor: Cultural Assimilation

- One of the biggest challenges to applications success is not technical, but rather cultural.



Critical Success Factor: caBIG on *SPEED*

- We have an arsenal of tools and capabilities that are already caBIG-compatible or can be made caBIG-compatible quickly.
  - These should be implemented as fast as possible.
  - They may not be the long term solution but they can stop the bleeding in certain areas.
  - If there are multiple choices, the focus should be on what brings the greatest value to the greatest number of users.

Critical Success Factor: Creating Value

- Value is simply the total future benefits to be realized by caBIG participants in terms of measurable results such as increase in productivity, decrease in cycle times or decrease in risk.
- Value measurements will vary from Workspace to Workspace, but a means for measuring how much value is being created is critical.

Value Creation in caBIG

- Think strategically, act tactically.
- Deliver quickly in priority order.
- Leverage the strength of existing systems and vendors.
- Unremitting focus on value. This includes:
  - Setting goals up front and always keeping those goals in the forefront as projects move forward.
  - Measuring results and adjusting if goals are not meeting expectations.
- Longer term, deliver “killer applications”
  - Applications that alter our approaches dramatically.
  - Applications that use technology to break the rules, implicit or explicit, dictating how cancer research is done.

Supporting the Mission

… of improving the cancer research enterprise and the lives of patients with cancer.